



PCT10

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/030,306

DATE: 08/06/2002
TIME: 12:45:43

Input Set : A:\661926.txt
Output Set: N:\CRF3\08062002\J030306.raw

Does Not Comply
Corrected Diskette Needed

3 <110> APPLICANT: Sagami Chemical Research Center,
4 Protegene Inc.
6 <120> TITLE OF INVENTION: Human proteins having hydrophobic domains and DNAs encoding
these
7 proteins
9 <130> FILE REFERENCE: 661926
C--> 11 <140> CURRENT APPLICATION NUMBER: US/10/030,306
C--> 11 <141> CURRENT FILING DATE: 2002-06-27
11 <150> PRIOR APPLICATION NUMBER: JP 11-194359
12 <151> PRIOR FILING DATE: 1999-07-08
14 <160> NUMBER OF SEQ ID NOS: 30

ERRORED SEQUENCES

1068 <210> SEQ ID NO: 25
1069 <211> LENGTH: 2150
1070 <212> TYPE: DNA
1071 <213> ORGANISM: Homo sapiens
1073 <220> FEATURE:
1074 <221> NAME/KEY: CDS
1075 <222> LOCATION: (32)...(1282)
1077 <400> SEQUENCE: 25
1078 ggtttctgcg ggtgaggctg gcgcccgtag c atg agc gag gcg gac ggg ctg 52
1079 Met Ser Glu Ala Asp Gly Leu 5
1080 1
1081 cga cag cgc cgg ccc ctg cgg ccg cag gtc gtc aca gac gat gat ggc 100
1082 Arg Gln Arg Arg Pro Leu Arg Pro Gln Val Val Thr Asp Asp Asp Gly 20
1083 10 15
1084 cag gcc ccg gag gct aag gac ggc agc tcc ttt agc ggc aga gtt ttc 148
1085 Gln Ala Pro Glu Ala Lys Asp Gly Ser Ser Phe Ser Gly Arg Val Phe 35
1086 25 30
1087 cga gtg acc ttc ttg atg ctg gct gtt tct ctc acc gtt ccc ctg ctt 196
1088 Arg Val Thr Phe Leu Met Leu Ala Val Ser Leu Thr Val Pro Leu Leu 55
1089 40 45 50
1090 gga gcc atg atg ctg ctg gaa tct cct ata gat cca cag cct ctc agc 244
1091 Gly Ala Met Met Leu Leu Glu Ser Pro Ile Asp Pro Gln Pro Leu Ser 70
1092 60 65
1093 ttc aaa gaa ccc ccg ctc ttg ctt ggt gtt ctg cat cca aat acg aag 292
1094 Phe Lys Glu Pro Pro Leu Leu Leu Gly Val Leu His Pro Asn Thr Lys 85
1095 75 80
1096 ctg cga cag gca gaa agg ctg ttt gaa aat caa ctt gtt gga ccg gag 340
1097 Leu Arg Gln Ala Glu Arg Leu Phe Glu Asn Gln Leu Val Gly Pro Glu 95
1098 90 100

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/030,306

DATE: 08/06/2002

TIME: 12:45:44

Input Set : A:\661926.txt

Output Set: N:\CRF3\08062002\J030306.raw

1099	tcc ata gca cat att ggg gat gtg atg ttt act ggg aca gca gat ggc	388
1100	Ser Ile Ala His Ile Gly Asp Val Met Phe Thr Gly Thr Ala Asp Gly	
1101	105 110 115	436
1102	cgg gtc gta aaa ctt gaa aat ggt gaa ata gag acc att gcc cgg ttt	
1103	Arg Val Val Lys Leu Glu Asn Gly Glu Ile Glu Thr Ile Ala Arg Phe	
1104	120 125 130 135	484
1105	ggt tgg ggc cct tgc aaa acc cga gat gat gag cct gtg tgt ggg aga	
1106	Gly Ser Gly Pro Cys Lys Thr Arg Asp Asp Glu Pro Val Cys Gly Arg	
1107	140 145 150	532
1108	ccc ctg ggt atc cgt gca ggg ccc aat ggg act ctc ttt gtg gcc gat	
1109	Pro Leu Gly Ile Arg Ala Gly Pro Asn Gly Thr Leu Phe Val Ala Asp	
1110	155 160 165	580
1111	gca tac aag gga cta ttt gaa gta aat ccc tgg aaa cgt gaa qtg aaa	
1112	Ala Tyr Lys Gly Leu Phe Glu Val Asn Pro Trp Lys Arg Glu Val Lys	
1113	170 175 180	628
1114	ctg ctg ctg tcc tcc gag aca ccc att gag ggg aag aac atg tcc ttt	
1115	Leu Leu Leu Ser Ser Glu Thr Pro Ile Glu Gly Lys Asn Met Ser Phe	
1116	185 190 195	676
1117	gtg aat gat ctt aca gtc act cag gat ggg agg aag att tat ttc acc	
1118	Val Asn Asp Leu Thr Val Thr Gln Asp Gly Arg Lys Ile Tyr Phe Thr	
1119	200 205 210 215	724
1120	gat tct agc agc aaa tgg caa aga cga gac tac ctg ctt ctg gtg atg	
1121	Asp Ser Ser Ser Lys Trp Gln Arg Arg Asp Tyr Leu Leu Val Met	
1122	220 225 230	772
1123	gag ggc aca gat gac ggg cgc ctg ctg gag tat gat act gtg acc agg	
1124	Glu Gly Thr Asp Asp Gly Arg Leu Leu Glu Tyr Asp Thr Val Thr Arg	
1125	235 240 245	820
1126	gaa gta aaa gtt tta ttg gac cag ctg cgg ttc ccg aat gga gtc cag	
1127	Glu Val Lys Val Leu Leu Asp Gln Leu Arg Phe Pro Asn Gly Val Gln	
1128	250 255 260	868
1129	ctg tct cct gca gaa gac ttt gtc ctg gtg gca gaa aca acc atg gcc	
1130	Leu Ser Pro Ala Glu Asp Phe Val Leu Val Ala Glu Thr Thr Met Ala	
1131	265 270 275	916
1132	agg ata cga aga gtc tac gtt tct ggc ctg atg aag ggc ggg gct gat	
1133	Arg Ile Arg Arg Val Tyr Val Ser Gly Leu Met Lys Gly Gly Ala Asp	
1134	280 285 290 295	964
1135	ctg ttt gtg gag aac atg cct gga ttt cca gac aac atc cgg ccc agc	
1136	Leu Phe Val Glu Asn Met Pro Gly Phe Pro Asp Asn Ile Arg Pro Ser	
1137	300 305 310	1012
1138	agc tct ggg ggg tac tgg gtg ggc atg tgg acc atc cgc cct aac cct	
1139	Ser Ser Gly Gly Tyr Trp Val Gly Met Ser Thr Ile Arg Pro Asn Pro	
1140	315 320 325	1060
1141	ggg ttt tcc atg ctg gat ttc tta tct gag aga ccc tgg att aaa agg	
1142	Gly Phe Ser Met Leu Asp Phe Leu Ser Glu Arg Pro Trp Ile Lys Arg	
1143	330 335 340	1108
1144	atg att ttt aag ctc ttt agt caa gag acg gtg atg aag ttt gtg ccg	
1145	Met Ile Phe Lys Leu Phe Ser Gln Glu Thr Val Met Lys Phe Val Pro	
1146	345 350 355	1156
1147	cgg tac agc ctc gtc cta gaa ctc agc gac agc ggt gcc ttc cgg aga	

RAW SEQUENCE LISTING PATENT APPLICATION: US/10/030,306

DATE: 08/06/2002
TIME: 12:45:44

Input Set : A:\661926.txt
Output Set: N:\CRF3\08062002\J030306.raw

1148 Arg Tyr Ser Leu Val Leu Glu Leu Ser Asp Ser Gly Ala Phe Arg Arg
1149 360 365 370 375 1204
1150 agc ctg cat gat ccc gat ggg ctg gtg gcc acc tac atc agc gag gtg
1151 Ser Leu His Asp Pro Asp Gly Leu Val Ala Thr Tyr Ile Ser Glu Val
1152 380 385 390 1252
1153 cac gaa cac gat ggg cac ctg tac ctg ggc tct ttc agg tcc ccc ttc
1154 His Glu His Asp Gly His Leu Tyr Leu Gly Ser Phe Arg Ser Pro Phe
1155 395 400 405 1300
1156 ctc tgc aga ctc agc ctc cag gct gtt tagccctccc agatagctgc c
1157 Leu Cys Arg Leu Ser Leu Gln Ala Val
1158 410 415 1360
1159 cctgccacgc aggccaggag tcttcacact caggcaccag gcctggtcca ggaggagctg 1420
1160 tggacacagt cgtggttcaa gtgtccacat gcacctgtta gtccctgaga ggtggtggga 1480
1161 atggctgctt cattcctcga ggatgccgg gccccacctg ggctgtctt tctgtttaga 1550 1570
E--> 1162 gggaagtgtg acatatctgc catgaggaac ataaattcat gtaaagccat tttctcttaa 1600
1163 acaaaacaaa actttctaag tacagtcatt ctctaggatt tgggaagctc cttgcacttg 1660
1164 gaacagggct caggtgggtg gagcagtaag gcactaccca gagagcttgc tgcgcggcc 1720
1165 ctgtcctgcg gcctcaaagt tcttctttac tatatataac gtgcggctcat acctttcttc 1780
1166 gttgtggtgg ggaatggaaga gcagagggag catggcccag ggggtgttgag gccagcggtg 1840
1167 agagccgtgt tagccaagac atggaactgt gttctcaagg gttatgtgg gcgtgggctc 1900
1168 tccatagtgt gtatgaaaag cttgttgact ctacggctc agagaggact ttgctgggtt 1960
1169 tctttctgtg aatatctccg tgcagccat gctggaattg gatgattctg caattcggga 2020
1170 cctaactgcag gggtcggtt agtaacgtct tgtctgtgat cttgttctt gacctctaga 2080
1171 cccaagatg tgaacagtgc acgtgttaat gtcactttg ctcattgtgtt ataagcccca 2140
1172 agttgctgta tattttcaca agtatgtcta cacactggtc atgattttga taataataa 2150
1173 cgataaatcg
1481 <210> SEQ ID NO: 30
1482 <211> LENGTH: 1892
1483 <212> TYPE: DNA
1484 <213> ORGANISM: Homo sapiens
1486 <220> FEATURE:
1487 <221> NAME/KEY: CDS
1488 <222> LOCATION: (5)...(1636)
1490 <400> SEQUENCE: 30
1491 agag atg gca gtg agc gag agg agg ggg ctc ggc cgc ggg agc ccc gcg 49
1492 Met Ala Val Ser Glu Arg Arg Gly Leu Gly Arg Gly Ser Pro Ala 15
1493 1 5 10 97
1494 gag tgg ggg cag cgg cta ctt ctg gtg ctg ctg ttg ggt ggc tgc tcc
1495 Glu Trp Gly Gln Arg Leu Leu Leu Val Leu Leu Leu Gly Gly Cys Ser 30
1496 20 25 145
1497 ggg cgc atc cac cgg ctg gcg ctg acg ggg gag aag cga gcg gac atc
1498 Gly Arg Ile His Arg Leu Ala Leu Thr Gly Glu Lys Arg Ala Asp Ile 45
1499 35 40 193
1500 cag ctg aac agc ttc ggt ttc tac acc aat ggc tct ctg gag gtg gag
1501 Gln Leu Asn Ser Phe Gly Phe Tyr Thr Asn Gly Ser Leu Glu Val Glu 60
1502 50 55 241
1503 ttg agc gtc ctg cgg ctg ggc ctc cgg gag gca gaa gag aag tcc ctg
1504 Leu Ser Val Leu Arg Leu Gly Leu Arg Glu Ala Glu Glu Lys Ser Leu
1505 65 70 75

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/030,306

DATE: 08/06/2002

TIME: 12:45:44

Input Set : A:\661926.txt

Output Set: N:\CRF3\08062002\J030306.raw

1506	ctg	gtg	ggg	ttc	agt	ctc	agc	cgg	gtt	cgg	tct	ggc	aga	gtt	cgc	tcc	289
1507	Leu	Val	Gly	Phe	Ser	Leu	Ser	Arg	Val	Arg	Ser	Gly	Arg	Val	Arg	Ser	95
1508	80					85				90							337
1509	tat	tca	acc	cgg	gat	ttc	cag	gac	tgc	cct	ctc	cag	aaa	aac	agt	agc	
1510	Tyr	Ser	Thr	Arg	Asp	Phe	Gln	Asp	Cys	Pro	Leu	Gln	Lys	Asn	Ser	Ser	110
1511				100					105								385
1512	agt	ttc	ctg	gtc	ctg	ttc	ctc	atc	aac	acc	aag	gat	ctg	cag	gtc	cag	
1513	Ser	Phe	Leu	Val	Leu	Phe	Leu	Ile	Asn	Thr	Lys	Asp	Leu	Gln	Val	Gln	125
1514				115					120								433
1515	gtg	cgg	aag	tat	gga	gag	cag	aag	acg	ttg	ttt	atc	ttt	ccc	ggg	ctc	
1516	Val	Arg	Lys	Tyr	Gly	Glu	Gln	Lys	Thr	Leu	Phe	Ile	Phe	Pro	Gly	Leu	140
1517				130					135								481
1518	ctc	ccg	gaa	gca	ccc	tcc	aaa	cca	ggg	ctc	ccg	aag	cca	cag	gcc	aca	
1519	Leu	Pro	Glu	Ala	Pro	Ser	Lys	Pro	Gly	Leu	Pro	Lys	Pro	Gln	Ala	Thr	
1520				145					150								529
1521	gtc	ccc	cgc	aag	gtg	gat	ggc	gga	ggg	acc	tct	gca	gcc	agc	aag	ccc	
1522	Val	Pro	Arg	Lys	Val	Asp	Gly	Gly	Gly	Thr	Ser	Ala	Ala	Ser	Lys	Pro	175
1523	160					165				170							577
1524	aag	tca	aca	ccc	gca	gtg	att	cag	ggt	cct	agt	ggg	aag	gac	aag	gac	
1525	Lys	Ser	Thr	Pro	Ala	Val	Ile	Gln	Gly	Pro	Ser	Gly	Lys	Asp	Lys	Asp	190
1526				180						185							625
1527	ctg	gtg	ttg	ggc	ctg	agc	cac	ctc	aac	aac	tcc	tac	aac	ttc	agt	ttc	
1528	Leu	Val	Leu	Gly	Leu	Ser	His	Leu	Asn	Asn	Ser	Tyr	Asn	Phe	Ser	Phe	205
1529				195					200								673
1530	cac	gtg	gtg	atc	ggc	tct	cag	gcg	gaa	gaa	ggc	cag	tac	agc	ctg	aac	
1531	His	Val	Val	Ile	Gly	Ser	Gln	Ala	Glu	Glu	Gly	Gln	Tyr	Ser	Leu	Asn	
1532				210					215								721
1533	ttc	cac	aac	tgc	aac	aat	tca	gtg	cca	gga	aag	gag	cat	cca	ttc	gac	
1534	Phe	His	Asn	Cys	Asn	Asn	Ser	Val	Pro	Gly	Lys	Glu	His	Pro	Phe	Asp	
1535				225					230								769
1536	atc	acg	gtg	atg	atc	cgg	gag	aag	aac	ccc	gat	ggc	ttc	ctg	tcg	gca	
1537	Ile	Thr	Val	Met	Ile	Arg	Glu	Lys	Asn	Pro	Asp	Gly	Phe	Leu	Ser	Ala	255
1538	240					245				250							817
1539	gcg	gag	atg	ccc	ctt	ttc	aag	ctc	tac	atg	gtc	atg	tcc	gcc	tgc	ttc	
1540	Ala	Glu	Met	Pro	Leu	Phe	Lys	Leu	Tyr	Met	Val	Met	Ser	Ala	Cys	Phe	270
1541				260					265								865
1542	ctg	gcc	gct	ggc	atc	ttc	tgg	gtg	tcc	atc	ctc	tgc	agg	aac	acg	tac	
1543	Leu	Ala	Ala	Gly	Ile	Phe	Trp	Val	Ser	Ile	Leu	Cys	Arg	Asn	Thr	Tyr	285
1544				275					280								913
1545	agc	gtc	ttc	aag	atc	cac	tgg	ctc	atg	gcg	gcc	ttg	gcc	ttc	acc	aag	
1546	Ser	Val	Phe	Lys	Ile	His	Trp	Leu	Met	Ala	Ala	Leu	Ala	Phe	Thr	Lys	300
1547				290					295								961
1548	agc	atc	tct	ctc	ctc	ttc	cac	agc	atc	aac	tac	tac	ttc	atc	aac	agc	
1549	Ser	Ile	Ser	Leu	Leu	Phe	His	Ser	Ile	Asn	Tyr	Tyr	Phe	Ile	Asn	Ser	315
1550				305					310								1009
1551	cag	ggc	cac	ccc	atc	gaa	ggc	ctt	gcc	gtc	atg	tac	tac	atc	gca	cac	
1552	Gln	Gly	His	Pro	Ile	Glu	Gly	Leu	Ala	Val	Met	Tyr	Tyr	Ile	Ala	His	335
1553	320					325				330							1057
1554	ctg	ctg	aag	ggc	gcc	ctc	ctc	ttc	atc	acc	atc	gcc	ctg	att	ggc	tca	

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/030,306

DATE: 08/06/2002
TIME: 12:45:44

Input Set : A:\661926.txt

Output Set: N:\CRF3\08062002\J030306.raw

1555 Leu Leu Lys Gly Ala Leu Leu Phe Ile Thr Ile Ala Leu Ile Gly Ser 350
 1556 340 345 1105
 1557 ggc tgg gcc ttc atc aag tac gtc ctg tcg gat aag gag aag aag gtc
 1558 Gly Trp Ala Phe Ile Lys Tyr Val Leu Ser Asp Lys Glu Lys Lys Val 365
 1559 355 360 1153
 1560 ttt ggg atc gtg atc ccc atg cag gtc ctg gcc aac gtg gcc tac atc
 1561 Phe Gly Ile Val Ile Pro Met Gln Val Leu Ala Asn Val Ala Tyr Ile 380
 1562 370 375 1201
 1563 atc atc gag tcc cgc gag gaa ggc gcc agc gac tac gtg ctg tgg aag
 1564 Ile Ile Glu Ser Arg Glu Gly Ala Ser Asp Tyr Val Leu Trp Lys 395
 1565 385 390 1249
 1566 gag att ttg ttc ctg gtg gac ctc atc tgc tgt ggt gcc atc ctg ttc
 1567 Glu Ile Leu Phe Leu Val Asp Leu Ile Cys Cys Gly Ala Ile Leu Phe 415
 1568 400 405 1297
 1569 ccc gta gtc tgg tcc atc cgg cat ctc cag gat gcg tct ggc aca gac
 1570 Pro Val Val Trp Ser Ile Arg His Leu Gln Asp Ala Ser Gly Thr Asp 430
 1571 420 425 1345
 1572 ggg aag gtg gca gtg aac ctg gcc aag ctg aag ctg ttc cgg cat tac
 1573 Gly Lys Val Ala Val Asn Leu Ala Lys Leu Lys Leu Phe Arg His Tyr 445
 1574 435 440 1393
 1575 tat gtc atg gtc atc tgc tac gtc tac ttc acc cgc atc atc gcc atc
 1576 Tyr Val Met Val Ile Cys Tyr Val Tyr Phe Thr Arg Ile Ile Ala Ile 460
 1577 450 455 1441
 1578 ctg ctg cag gtg gct gtg ccc ttt cag tgg cag tgg ctg tac cag ctc
 1579 Leu Leu Gln Val Ala Val Pro Phe Gln Trp Gln Trp Leu Tyr Gln Leu 475
 1580 465 470 1489
 1581 ttg gtg gag ggc tcc acc ctg gcc ttc ttc gtg ctc acg ggc tac aag
 1582 Leu Val Glu Gly Ser Thr Leu Ala Phe Phe Val Leu Thr Gly Tyr Lys 495
 1583 480 485 1537
 1584 ttc cag ccc aca ggg aac ccg tac ctg cag ctg ccc cag gag gac
 1585 Phe Gln Pro Thr Gly Asn Asn Pro Tyr Leu Gln Leu Pro Gln Glu Asp 510
 1586 500 505 1585
 1587 gag gag gat gtt cag atg gag caa gta atg acg gac tct ggg ttc cgg
 1588 Glu Glu Asp Val Gln Met Glu Gln Val Met Thr Asp Ser Gly Phe Arg 525
 1589 515 520 1633
 1590 gaa ggc ctc tcc aaa gtc aac aaa aca gcc agc ggg cgg gaa ctg tta
 1591 Glu Gly Leu Ser Lys Val Asn Lys Thr Ala Ser Gly Arg Glu Leu Leu 540
 1592 530 535 1690
 1593 tgateac ctccacatct cagaccaaag ggtcgctctc ccccgacatt tctcactct 1750
 1594 gcccttcttc cacagcgtat gtggggaggt ggagggggtc catgtggacc aggcgcccag 1810
 1595 ctccccggga ccccggttcc cggacaagcc catttggaag aagagtcct tctcccccc 1870
 1596 aaatattggg cagccctgtc cttaccccg gaccacccct ccttcacagc tatgtgtaca 1892
 1597 ataattgacca atctgtttgg ct

E--> 1599 1
 E--> 1601 1
 E--> 1602 1

/64

Lolite

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/030,306

DATE: 08/06/2002
TIME: 12:45:45

Input Set : A:\661926.txt

Output Set: N:\CRF3\08062002\J030306.raw

L:11 M:270 C: Current Application Number differs, Replaced Current Application No
L:11 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:1162 M:254 E: No. of Bases conflict, LENGTH:Input:1550 Counted:1540 SEQ:25
L:1599 M:254 E: No. of Bases conflict, this line has no nucleotides.
M:254 Repeated in SeqNo=30
L:1602 M:320 E: (1) Wrong Nucleic Acid Designator, NUMBER OF INVALID KEYS:1
L:1602 M:252 E: No. of Seq. differs, <211> LENGTH:Input:1892 Found:1893 SEQ:30